

SCORE Search Results Details for Application 10552515 and Search Result 20080624_083152_us-10-552-515-1.rpr.

Score Home	Retrieve Application	SCORE System	SCORE	Comments /
Page	List	Overview	FAQ	Suggestions

This page gives you Search Results detail for the Application 10552515 and Search Result 20080624_083152_us-10-552-515-1.rpr.

[Go Back to previous page](#)

GenCore version 6.2.1

Copyright (c) 1993 - 2008 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 24, 2008, 08:36:37 ; Search time 36 Seconds
(without alignments)
2493.618 Million cell updates/sec

Title: US-10-552-515-1
Perfect score: 4950
Sequence: 1 MRMAATAWAGLQGPPLPTLC.....SELSSHWTPTVTPKASQLQQ 933

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_80:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	734	14.8	1049	2	T22762	hypothetical prote
2	288.5	5.8	572	2	F96755	hypothetical prote
3	181.5	3.7	946	2	S48255	probable membrane
4	117	2.4	548	2	I48693	natural resistance
5	115.5	2.3	3010	1	GNWVTC	genome polyprotein
6	110.5	2.2	680	2	T35404	probable squalene-
7	110.5	2.2	873	2	S46584	probable membrane
8	110	2.2	792	2	T00487	probable potassium

9	108	2.2	3010	1	A45573	genome polyprotein
10	106.5	2.2	519	2	T11129	cytochrome-c oxida
11	106	2.1	438	2	B86088	probable citrate p
12	106	2.1	438	2	E91240	probable membrane
13	105	2.1	621	2	JC1346	dopamine beta-mono
14	104	2.1	646	2	H82555	c-type cytochrome
15	103.5	2.1	478	2	JQ2034	RNA-directed RNA p
16	102	2.1	302	2	C83993	hypothetical prote
17	101.5	2.1	395	2	D81040	cytochrome c-type
18	101.5	2.1	395	2	B81986	probable membrane
19	101	2.0	466	2	A95355	probable inner-mem
20	100.5	2.0	585	2	S74673	pleD protein - Syn
21	100	2.0	515	2	D71390	cytochrome-c oxida
22	100	2.0	3010	1	GNWVCJ	genome polyprotein
23	99.5	2.0	737	2	AG2156	hypothetical prote
24	98.5	2.0	413	2	AF0393	NADH2 dehydrogenas
25	98.5	2.0	1353	2	T26301	hypothetical prote
26	98.5	2.0	1755	2	S69845	TyB protein - yeas
27	98	2.0	348	2	T12280	NADH2 dehydrogenas
28	98	2.0	1265	2	T51314	probable CO-induce
29	97.5	2.0	348	2	T12291	NADH2 dehydrogenas
30	97.5	2.0	348	2	T12290	NADH2 dehydrogenas
31	97.5	2.0	460	2	A84154	amino acid transpo
32	97.5	2.0	906	2	G83156	probable transcrip
33	96.5	1.9	348	2	T12281	NADH2 dehydrogenas
34	96.5	1.9	417	2	C81084	probable integral
35	96.5	1.9	491	2	B70414	NADH2 dehydrogenas
36	96.5	1.9	1755	2	S69969	TyB protein - yeas
37	96	1.9	419	1	SYPJCD	naringenin-chalcon
38	96	1.9	428	2	T48284	hypothetical prote
39	96	1.9	865	2	T40288	hypothetical prote
40	95.5	1.9	572	2	T48601	hypothetical prote
41	95.5	1.9	758	2	D71072	hypothetical prote
42	95.5	1.9	1755	2	S50663	TyB protein - yeas
43	95	1.9	429	2	AG3150	hypothetical prote
44	95	1.9	473	2	AC0479	glycerol-3-phospha
45	95	1.9	631	2	B98137	hypothetical 46.1K

ALIGNMENTS

RESULT 1
T22762
hypothetical protein F56A8.1 - *Caenorhabditis elegans*
C;Species: *Caenorhabditis elegans*
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T22762
R;McMurray, A.
submitted to the EMBL Data Library, December 1996
A;Reference number: Z19612
A;Accession: T22762
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1049 <WIL>
A;Cross-references: UNIPROT:O45572; UNIPARC:UPI000007F44C; EMBL:Z83230; PIDN:CAB05741.1; GSPDB:GN00021; CESP:F56A8.1
A;Experimental source: clone F56A8
C;Genetics:
A;Gene: CESP:F56A8.1
A;Map position: 3
A;Introns: 86/3; 146/3; 208/3; 245/2; 295/2; 325/2; 397/3; 532/3; 582/2; 612/3; 654/1; 677/1; 707/3; 734/3; 786/2; 812/2; 870/1; 902/3; 942/3; 1011/2

http://es/ScoreAccessWeb/GetItem.action?AppId=105525...24_083152_us-10-552-515-1.rpr&ItemType=4&startByte=0 (3 of 20)10/10/2008 8:49:35 AM

```

      ||  ::| || |:: :::: |          :|| | :| |: : :
Db      716 LLSSVIP SIPASIKLALRKKRYVVAHIVEKGDVPHRTRIKKRTRIAKLAWIASNQKMIKS 775

Qy      904 NGTKDEQPKGSELSSHWT 921
      | |:: | |::|
Db      776 NRKKEKSNK-----KHFT 788
```

RESULT 2

F96755

hypothetical protein F3N23.22 [imported] - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004

C;Accession: F96755

R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, J.; Altaf, H.; Araujo, R.; Bowman, C.L.; Brooks, S.Y.; Buehler, E.; Chan, A.; Chao, Q.; Chen, H.; Cheuk, R.F.; Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; Dunn, P.; Etgu, P.; Feldblyum, T.V.; Feng, J.; Fong, B.; Fujii, C.Y.; Gill, J.E.; Goldsmith, A.D.; Haas, B.; Hansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.J.; Koo, H. L.; Kremenetskaia, I.; Kurtz, D.B.; Kwan, A.; Lam, B.; Langin-Hooper, S.; Lee, A.; Lee, J.M.; Lenz, C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali, A.; Militischer, J.; Miranda, M.; Nguyen, M.; Nierman, W.C.; Osborne, B.I.; Pai, G.; Peterson, J.; Pham, P.K.; Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, L.J.; Tambunga, G.; Toriumi, M.J.; Town, C.D.; Utterback, T.; van Aken, S.; Vaysberg, M.; Vysotskaia, V. S.; Walker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A;Reference number: A86141; MUID:21016719; PMID:11130712

A;Accession: F96755

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-572 <STO>

A;Cross-references: UNIPROT:Q9SSM5; UNIPARC:UPI00000A63BD; GB:AE005173; NID:g5903091; PIDN:

AAD55649.1; GSPDB:GN00141

C;Genetics:

A;Gene: F3N23.22

A;Map position: 1

```

Query Match          5.8%;  Score 288.5;  DB 2;  Length 572;
Best Local Similarity 20.5%;  Pred. No. 3.4e-16;
Matches 169;  Conservative 97;  Mismatches 236;  Indels 323;  Gaps 29;
```

```

Qy      142 ETFLDNLRAAGLCVDQQDVQDGNTTVHYALLSASWAVLCYYAEDLRLKLPLQELPNQASN 201
      | : || |: ||:          |:  | |:: | : | | |:
Db      32 EVLVTELRRKGMVVDL-----VVGLAHEFLKVAAPSEILGNAAAE 71

Qy      202 WSAGLLAWLGIPNVLLLEVVPDPPEYYSCRFRVKNLPRFLGSDNQDTFFTSTKRHQILFE 261
      |||          |:| | : :  | | :| : :| |
Db      72 LHIRKPTRLGI-----DLPFEMQGSEAFIRQPDGLLFS-----WFERFRCYQHLY 117

Qy      262 ILAKTPYGHEKKNLLG-----IHQLLAEGVLSAAFPLHDGPFKTPPEGPQAPR 309
      : : ||: | : : | : ||: | |||
Db      118 GIVNSG-GHDVTLKLDGREFCWTAGESLLRRLESEGVKQMFPLHDE----- 163

Qy      310 LNQRQVLFQHWA-RWGKWN-KYQPLDHVRRYFGEKVALYFAWLGFYTGWLLPAAVVGTLV 367
      |: | |:|| | ||  |:| : ||| |
Db      164 -LKRKELLQNWALNW--WNCTNQPIDQIYSYFGAK----- 195

Qy      368 FLVGCFLVFSDIPTQELCGSKDSFEMCPLCLDCPFWLLSSACALAQAGRLFDHGGTVFFS 427
      | : |
Db      196 -----ELIKNLGN----- 203
```

Qy	428	LFMALWAVLLLEYWKRKSATLAYRWDCSDYEDTEERPRPQFAASAPMTAPNPITGEDEPY	487
Db	204	-----ERAKEKEAYQRYEW-----	217
Qy	488	FPERSRARRMLAGSVVIVVMVAVVVMCLVSIILYRAIMAIVVSRSGNTLLAAWASRIASL	547
		: : : : : : :	
Db	218	FAYRKRFNR-----DVLVIMSIIICLQLPFELAYAHIFEIITSDI IKYVLTA-----	263
Qy	548	TGSVVNLVFILILSKIYVSLAHVLTWRWEMHRTQTKFEDAFTLKVFIFQFVNFYSSPVYIA	607
		: : : : : : : : :	
Db	264	----IYLLIIQYLTRLGGKVSVKLINREINESVEYRANSLIYKVF-----GLYFMQTYIG	314
Qy	608	FFKGRFVGYPGNYHTLFGVRNEECAAGGCLIELAQELLVIMVGKQVINNMQEVLI PKLKG	667
		: : : : : :	
Db	315	IF-----YHVLLH-RN-----FMTLRQVLIQRLIISQVFWTLMDGSLPYLKY	355
Qy	668	WWQKFRRLRSKKR-KAGASAGASQ--GPWEDDY-----ELVPCEGLFDEYLEMVL	713
		: : : : : : : :	
Db	356	SYRKYRARTKKKMEDGSSTGKIQIASRVEKEYFKPTYASIGVELE--DGLFDDSLELAL	413
Qy	714	QFGFVTIFVAACPLAPLFALLNNWVEIRLDARKFVCEYRRPVAERAQDIGIWFHILAGLT	773
		: : : : : : : :	
Db	414	QFGMIMMFACAFPLAFALAAVSNVMEIRTNALKLLVTLRRPLPRAAATIGAWLNIWQFLV	473
Qy	774	HLAVISNAFLLAFFSSDFLPRAYYRWTRAHDLRGFLNFTLARAPSSFAAAHNRTCryRAFR	833
		: : : : :	
Db	474	VMSICTNSALL-----VCLY-----	488
Qy	834	DDDGHYSTYWNLLAIRLAFVIVFEHVVSFVGRLLDLLVPDIPESVEI-KVK-----REY	887
		: : : : : : : : : : :	
Db	489	DQEGKWK-----IEPGLAAILIMEHVLLLLKFGLSRLVPEEPAWVRASRVKNVTQAQDM	542
Qy	888	YLAQKQALAENEVLFGTNGTKDEQPKGSELSSHWPFTVPKASQLQ	932
		: : :	
Db	543	Y-CKQLL-----RSISGEFNSLTKEPEQQQQ	567

RESULT 3
S48255
probable membrane protein YBR086c - yeast (*Saccharomyces cerevisiae*)
N;Alternate names: hypothetical protein YBR0809
C;Species: *Saccharomyces cerevisiae*
C;Date: 03-Aug-1995 #sequence_revision 11-Aug-1995 #text_change 09-Jul-2004
C;Accession: S48255; S45954; S44670
R;Mannhaupt, G.; Stucka, R.; Ehnle, S.; Vetter, I.; Feldmann, H.
Yeast 10, 1363-1381, 1994
A;Title: Analysis of a 70 kb region on the right arm of yeast chromosome II.
A;Reference number: S48255; MUID:95208357; PMID:7900426
A;Accession: S48255
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-946 <MAN>
A;Cross-references: UNIPROT:P38250; UNIPARC:UPI0000036C25; EMBL:X78993; NID:g476045; PIDN:CAA55593.1; PID:g476046
R;Feldmann, H.; Mannhaupt, G.; Schwarzlose, C.; Vetter, I.
submitted to the Protein Sequence Database, August 1994
A;Reference number: S45927
A;Accession: S45954
A;Molecule type: DNA
A;Residues: 1-946 <FE2>
A;Cross-references: UNIPARC:UPI0000036C25; EMBL:Z35955; NID:g536351; PID:g536352; MIPS:YBR086c
C;Genetics:

A;Gene: SGD:IST2
A;Cross-references: SGD:S0000290
A;Map position: 2R
C;Superfamily: Saccharomyces cerevisiae probable membrane protein YBR086c
C;Keywords: transmembrane protein
F;131-147/Domain: transmembrane #status predicted <TM1>
F;158-174/Domain: transmembrane #status predicted <TM2>
F;207-243/Domain: transmembrane #status predicted <TM3>
F;248-274/Domain: transmembrane #status predicted <TM4>
F;302-324/Domain: transmembrane #status predicted <TM5>
F;450-477/Domain: transmembrane #status predicted <TM6>
F;506-532/Domain: transmembrane #status predicted <TM7>
F;563-588/Domain: transmembrane #status predicted <TM8>

Query Match 3.7%; Score 181.5; DB 2; Length 946;
Best Local Similarity 18.6%; Pred. No. 8.6e-07;
Matches 118; Conservative 99; Mismatches 243; Indels 173; Gaps 25;

Qy	342	KVALYFAWLGFYTGWLLPAAVVGTLVFLVGCFLVFSDIPTQELCGSKDSFEMCPLCLDCP	401
		: : : : :	
Db	119	KQSLYFAFLQNYIKWLIPFSFFGLSIRFLSNF-----	150
Qy	402	FWLLSSACALAQAGRLFDHGGTVFFSLFMALWAVLLLEYWKRK-----SATLAYRWDCSD	456
		: : : : : :	
Db	151	-----TYEFNST--YSLFAILWTLSTAFWLYKYEPFWSDRLSKYSSFST	193
Qy	457	YEDTEERPRPQFAASAPMTAPN----PITGEDEPYFPERSRARRMLAGSVVIVVMVAVVV	512
		: : : : : : : : : : : :	
Db	194	IEFLQDKQKAQKKASSVIMLKCCFIPVA-----LLFGA----ILLSFQL	234
Qy	513	MCLVSIILYRAIMAIVVSRSNTLLAAWASRIASLTGSVNVLVFILILSKIYVS-LAHVL	571
		: : : : : : : : : :	
Db	235	YCFALIEIFYKQIY-----NGPMI-----SILSFLPTILICTFTPVLTVIYNKYFVEPM	282
Qy	572	TRWEMHRTQTKFEDAFTLKVFIFQFVNFYSSPVYIAFFKGRFVGYPGNYHTLFGVRNEEC	631
		: : : : : : : : : : :	
Db	283	TKWENHSSVVNNAKKSKEAKNFVIFLSSY-VPLLITL----FLYLPMGHLLTAEIRTKVF	337
Qy	632	AAGGCLIEL-----AQELLVIMVGKQVINNMQEVLPKLKGWWQK-----	671
		: : : : : : :	
Db	338	NAFSILARLPHTDSDFIIDTKRYEDQFFYFIVINQLIQFSMENFVPSLVSIAQQKINGPN	397
Qy	672	---FRLRSKKRKAGASAGASQGPWE--DDYELVPCEGLFD---EYLEMVLQFGFVTIFVA	723
		: : : : : : : : : : :	
Db	398	PNFVKAESEIGKAQLSS-SDMKIWSKVKSQYQTDWPWGATFDLDANFKKLLQLFGYLVMFST	456
Qy	724	ACPLAPLFPALLNNWVEIRLDARKFVC----EY-RRPVAERAQD-----IGIWFHILA	770
		: : : : : : : : :	
Db	457	IWPLAPFICLIVNLIVYQVDLRKAVLYSKPEYFPFPIYDKPSSVSNTQKLTVGLWNSVLV	516
Qy	771	GLTHL-AVISNAFLAFSSDFLPRAYRWTRAHDLRGFLNFTLARAPSSFAAAHNRTCry	829
		: : : : : : : : :	
Db	517	MFSILGCVITATLTMYQSCNIP-----GVGAHTSIHTNKAWY	554
Qy	830	RAFRRDDGHYSQTYWNLLAIRLAFVIVFEHVVSFVGRLLDLLVPDIPESVEIKVKREYYL	889
		: : : : : : : : : :	
Db	555	LA-----NPINHSWINI---VLYAVFIEHVSVAIFFLFSILKSSHDDVANGIVPKHV	605
Qy	890	AKQALAEENEVL-----FGTNGTKD-EQPKGS	914
		: : : : :	
Db	606	NVQNPPKQEVFEKIPSPEFNSNNEKELVQRKGS	638

Qy	67	VLIDVSPPEAEKRGSYGSTAHASEPG-GQQAAACRAGSPAKPRIADFVLVWEEDLKLDRO	125
		:: :	
Db	1	MISDKSPRL-SRPSYSGSI--SSLPGPAPQPAPCR-----ETYLSEKIP	41
Qy	126	QDSAARDRTDMHRTWRET-----FLD--NLRAAGLCVDQQQDVQDGNTTVHYALLSA	174
		: : : : :	
Db	42	IPSADQGTFSLRKLAFTGPGFLMSIAFLDPGNI-----ESDLQAGAVAGFKLLWL	93
Qy	175	SWA----VLC-----YYAEDLRKL-----PLQELPNQ	198
		: : : : :	
Db	94	LWATVLGLLCQRLAARLGVTGKDLGEVCHLYYPKVPRILLWTIELAIVGSMDQEVIQT	153
Qy	199	ASNW---SAGLLAWLGIPNVLLLEVVPDPVPEEYSCRFRVNKLPRFLGSDNQDTFFTSTKR	255
		:: : : : : :	
Db	154	AISFNLLSAGRIPLWG--GVLTITIV-DTFFFLFDNYGLRKLEAFFG-----	197
Qy	256	HQILFEILAKTPYGHEKKNLLGIHQLLAEGVLSAAFPLHDGPFKTPPEGPQAPRLNQRV	315
		: : : : : : : : :	

Db	198	--LLITIMALT-FGYE---YVVAHP--SQGALLKGLVLPTCPGCGQPELLQAVGIVGAI	249
Qy	316	LFQHWARWGKWNKYQPLDHVRRYFGEKVALYFAWLGFYTGWLLPAAVVGTILVFLVGCFL-	374
		: : : : : : : : :: :: :	
Db	250	MPHNIYLHSALVKSREVDRTTRRVDVREANMYF-----LIEATIALSVSFIINLFVM	300
Qy	375	-VFSDIPTQELCGSKDSFEMCPLCLDCPFWLLSSACALAQAGRLFD-----HGG	422
		: :::: : : : : :	
Db	301	AVFGQAFYQQT--NEEAFNIC-----ANSSLQNYAKIFPRDNNTVSVDIYQGG	346
Qy	423	TVFFSLF---MALWAVLLLEYWKRKSATLAY-----RWDCSDYEDTEERPRP	466
		: : : : :	
Db	347	VILGCLFGPAALYIWAAGQSSMTGTGTYAGQFVMEGFLKLRW-----	392
Qy	467	QFAASAPMTAPNPITGEDEPYFPERSR-ARRMLAGSVVIV--VMVAV-----V	511
		: : : :	
Db	393	-----SRFARVLLTRSCAILPTVLVAVFRDLKDLGLNDL	427
Qy	512	VMCLVSIILYRAIMAIVVSRSNTLLAAWAS-RIASLTGS-----VVNLVFILILSKI	563
		: :: : : : : : : : : : ::	
Db	428	LNVLQSLLLPFAVLPILTFTSMPAVMQEFANGRMSKAITSCIMALVCAINLYFVI---SY	484
Qy	564	YVSLAHVLTRWEMHRTQTKFEDAFTLVKFIFQFVNFYSSPVYIAF-----FKGRFVGYPG	618
		: : : : : : : :	
Db	485	LPSLPH-----PAYFGLVALFA-IGYLGLTAYLAWTCCIAHGATFLTHSS	528
Qy	619	NYHTLFGVRNEE	630
		: :	
Db	529	HKHFLYGLPNEE	540

RESULT 5
GNWVTC
genome polyprotein - hepatitis C virus
N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstructural protein NS3); major envelope protein E; nonstructural protein NS1; nonstructural protein NS2; nonstructural protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
C;Species: hepatitis C virus
C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 09-Jul-2004
C;Accession: A38465
R;Takamizawa, A.; Mori, C.; Fuke, I.; Manabe, S.; Murakami, S.; Fujita, J.; Onishi, E.; Andoh, T.; Yoshida, I.; Okayama, H.
J. Virol. 65, 1105-1113, 1991
A;Title: Structure and organization of the hepatitis C virus genome isolated from human carriers.
A;Reference number: A38465; MUID:91140698; PMID:1847440
A;Accession: A38465
A;Molecule type: genomic RNA
A;Residues: 1-3010 <TAK>
A;Cross-references: UNIPROT:P26663; UNIPARC:UPI0000131E1C; EMBL:M58335; NID:g329770; PIDN:AAA72945.1; PID:g329771
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: ATP; capsid protein; envelope protein; glycoprotein; hydrolase; nonstructural protein; nucleotide binding; P-loop; polyprotein; serine proteinase; transmembrane protein
F;2-115/Product: capsid protein C #status predicted <CPC>
F;116-191/Product: envelope protein M #status predicted <EPM>
F;192-389/Product: major envelope protein E #status predicted <MEE>
F;390-729/Product: nonstructural protein NS1 #status predicted <NS1>
F;730-1006/Product: nonstructural protein NS2 #status predicted <NS2>
F;1007-1615/Product: hepacivirin #status predicted <NS3>
F;1230-1237/Region: nucleotide-binding motif A (P-loop)
F;1312-1317/Region: nucleotide-binding motif B
F;1316-1319/Region: DEXH motif
F;1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>

F;1863-2013/Product: nonstructural protein NS4b #status predicted <N4B>
F;2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>
F;196,209,234,250,305,325,417,423,430,448,532,540,556,576,623,645,1213,1255,2041,2077,2240,2529,2788/
Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 2.3%; Score 115.5; DB 1; Length 3010;
Best Local Similarity 19.7%; Pred. No. 1.7;
Matches 163; Conservative 83; Mismatches 285; Indels 295; Gaps 40;

Qy	128	SAARDRTDMHRTWRETFLDNLRAAGLCVDQ-----QDVQDGNTTVHYALLSASWAVLCYY	182
		: :	
Db	314	SGHRMAWDMMMNWSPT-----TALVVSQLLRIPQAVVDMVAGAHWGVL----AGLAYY	362
Qy	183	AEDLRLKLPLQELPNQASNWSAGLLAWLGIPNVLLEVVPDPPEYYSCRFRVKNLPRFLG	242
		: : :	
Db	363	-----SMAGNWAKVLIVML-----LFAG	380
Qy	243	SDNQDTFFT-----STKRHQILF-----EILAKTPYGHEKKNLLGIHQLLAEGVLS	288
		: : : : : : :	
Db	381	VDG-DTHVTGGAQAKTTNRLVSMFASGPSQKIQLINTNGSWHINRTALNCNDSLQTGFLA	439
Qy	289	AAFPLHDGPFKTPPE-----GP-----QAPRLNQRQVLFQH-----	319
		: : : : :	
Db	440	ALFYTHSFNSSGCPERMAQCRTIDKFDQGWGPITYAESSRSDQRPYCWHYPPPOCTIVPA	499
Qy	320	-----WARWGK-----WNKYQPLDHVRRYFGEKVALY	346
		: : :	
Db	500	SEVCGPVYCFTPSPVVVGTTDRFGVPTYRWGENETDVLLLNNTRPPQ--GNWFG-----	551
Qy	347	FAWL---GFYTGWLLPAAVVG-----TLVFLVGCFLVFSDIPTQELCGSKDSFEMCPLCL	398
		: : : : :	
Db	552	CTWMNSTGFTKTCGGPPCNIGGVGNNTLTCTDCFRKHPE-ATYTKCGSGP--WLTPRCM	608
Qy	399	-DCPFWLLSSACALAQAGRLFDHGGTVFFSLFMALWAVLLLEYWKRKSATLAYRWDCSDY	457
		: : : : : : :	
Db	609	VDYPY-----RLWHYPCTVNFTIFKVRMYVGGVEH--RLNA--ACNWTGRGER	651
Qy	458	EDTEERPRPQFAASAPMTAPNPITGEDEPYFPERSRARRMLAGSVVIVVMVAVVVMCLVS	517
		: : : : : : : :	
Db	652	CDLEDRDRPELSPLLLSTTEWQVLPSCSFTTLPALSTGLIHLHQNIVDVQYLYGIGSAVVS	711
Qy	518	IIL---YRAIMAIVVSRSGNTLLAAWAS-RIASLTGSVNLVFIILILSKIYVSLAHLVTR	573
		: : : : : : : :	
Db	712	FAIKWEYVLLLFLLLA-DARVCACLWMMLLIAQAEAALENLV---VLNSASVAGAH----	763
Qy	574	WEMHRTQTKFEDAFTLKVFIFQFVNFYSSPVYIAFFKGRFVGYPGNYHTLFGV-----	626
		: : : : :	
Db	764	-----GILSFLVFFCAAWYI---KGRLV--PGATYALYGVWPLLLLLL	800
Qy	627	-----RNEECAAGGCLIELAQELLVIMVGKQV--INNMQEVLIPLKLGWWQKFR	673
		: : : : : : :	
Db	801	LALPPRAYAMDREMAASCGG-----AVFVGLVLLTSLSPYYKVFLARLIWWLQYFT	850
Qy	674	LRSKKRKAGASAGASQGPW-----EDDYELVPC----EGLFDEYLEMVLFQGFVTI	720
		: : : : : :	
Db	851	TR-----AEADLHVWIPPLNARGGRDAIILLMCAVHPELIFDITKLLIAILGPLMV	901
Qy	721	FVAACPLAPLFAALLNNWVEIRLDARKFVCEYRRPVAERAQDIGIWFHILAGLT-----H	774
		: : :	
Db	902	LQAGITRVPIYF-----VRAQGLIHACMLVRKVA--GGHYVQMAFMKLGALTGTIYINH	952
Qy	775	LAVISNAFLLAFSSDFLPRAYYRWTRAHDLRGFLNFTLARAPSSFA	820
		: : : : :	

Db 953 LTPLRD-----WPRA----GLRDLAVAVEPVPVFS 977

RESULT 6

T35404

probable squalene-hopene cyclase - Streptomyces coelicolor

C;Species: Streptomyces coelicolor

C;Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 09-Jul-2004

C;Accession: T35404

R;Oliver, K.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, March 1999

A;Reference number: Z21577

A;Accession: T35404

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-680 <OLI>

A;Cross-references: UNIPROT:Q9X7V9; UNIPARC:UPI00000DAF47; EMBL:AL049485; PIDN:CAB39697.1; GSPDB:
GN00070; SCOEDB:SC6A5.13

A;Experimental source: strain A3(2)

C;Genetics:

A;Gene: SCOEDB:SC6A5.13

C;Superfamily: squalene-hopene cyclase

Query Match 2.2%; Score 110.5; DB 2; Length 680;
Best Local Similarity 22.8%; Pred. No. 0.61;
Matches 112; Conservative 52; Mismatches 173; Indels 155; Gaps 27;

Qy 40 MTSETSSGSHCARSRMLRRRAQEEDSTVLIDVSPPEAEKRGSYGSTAHASEPGGQQAAC 99
||: |: || | | | : | |: | : | |
Db 1 MTA-TTDGSTGASLRPLAASASDITDI-----PAAAAGVPEAAA- 39

Qy 100 RAGSPAKPRIADFLV-----WEEDLKLD RQQDSAARDRTDMHRTWRETFL-----DN 147
| | ||:| : | | : : | |: | | :
Db 40 ----RATRRATDFLLAKQDAEGWWKGD-----ETNVTMDAEDL---LLRQFLGIQDEET 87

Qy 148 LRAAGLCVDQQDVQDGNNTTVHY-ALLSASWAVLCYYAEDLRKLPLQELPN--QASNW-- 202
||| | : : : || | : | | || | |: : |: |
Db 88 TRAAALFIRGEQREDGTWATFYGGPGELSTTIEAYVA--LRLAGDSPEAPHMARAAEWIR 145

Qy 203 SAGLLA-----WLGIPN-VLLEVVPDPVPE--YYSRFRVKNLPRFLGSDNQDTFFT 251
| |:| | | : : : ||| |: | : : | |
Db 146 SRGGIASARVFTRIWLALFGWWKWDLPPELPELIYF-----PTWVPLNIYD--FG 194

Qy 252 STKRHQI--LFEILAKTPYGHEKKNLLGIHQLLAEGVLSAAFP---LHDGPFKTPPEGPQ 306
| | | : || | | | | | | : | |
Db 195 CWARQTIVPLTIVSAKRP-----VRPAPFPLDELHTDPARPNNPPRPL 236

Qy 307 AP-----RLNQ-----RQVLFQHWARW-----GKWNKYQPLDHVR 336
|| |:: | : || | | | |
Db 237 APVASWDGAFQRIDKALHAYRKVAPRRLRAAMNSAARWIIERQENDGCWGGIQP----- 291

Qy 337 RYFGEKVALYFAWLGFYTGWLLPAAVVGTLVFLVGCFLVFSDIPTQELCGSKDSFEMCPL 396
|:| : |: | |: : : | |: : | |
Db 292 -----PAVYSVIALYLLGYDLEHPVMRAGLESIDRFVWRE-----DGARMIEA 335

Qy 397 CLDCPFW-LLSSACALAQAGRLFDHGGTVFFSLFMALWAVLLLEYWKRKSATL-----AY 450
| | | : ||| || | | | : : | | | :
Db 336 C-QSPVWDTCLATIALADAGVPEDHPQLVKASDWMLGEQIVRPGDWSVKRPGLPPEGWAF 394

Qy 451 RWDCSDYEDTEE 462
: : | | : :
Db 395 EFHNDNYPDIDD 406

RESULT 7

S46584

probable membrane protein YJL094c - yeast (*Saccharomyces cerevisiae*)

N;Alternate names: hypothetical protein J0909

C;Species: *Saccharomyces cerevisiae*

C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004

C;Accession: S46584; S56871; S47057

R;Miosga, T.; Witzel, A.; Zimmermann, F.K.

Yeast 10, 965-973, 1994

A;Title: Sequence and function analysis of a 9.46 kb fragment of *Saccharomyces cerevisiae* chromosome X.

A;Reference number: S46584; MUID:95076716; PMID:7985424

A;Accession: S46584

A;Molecule type: DNA

A;Residues: 1-873 <MIO>

A;Cross-references: UNIPROT:P40309; UNIPARC:UPI000013B5C4; EMBL:X77087; NID:g521093; PIDN:CAA54359.1; PID:g521094

A;Note: the authors translated the codon TCC for residue 645 as Trp

R;Miosga, T.; Schaaff-Gerstenschlaeger, I.; Baur, A.; Boles, E.; Chalwatzis, N.; Fournier, C.; Schmitt, S.; Velten, C.; Wilhelm, N.; Witzel, A.; Zimmermann, F.K.

submitted to the Protein Sequence Database, September 1995

A;Reference number: S56855

A;Accession: S56871

A;Molecule type: DNA

A;Residues: 1-873 <MIW>

A;Cross-references: UNIPARC:UPI000013B5C4; EMBL:Z49369; NID:g1008267; PID:g1008268; MIPS:YJL094c

C;Genetics:

A;Gene: SGD:KHA1

A;Cross-references: SGD:S0003630

A;Map position: 10L

C;Keywords: transmembrane protein

Query Match 2.2%; Score 110.5; DB 2; Length 873;

Best Local Similarity 21.4%; Pred. No. 0.86;

Matches 69; Conservative 56; Mismatches 106; Indels 91; Gaps 16;

Qy	503	VIVVMVAVVVMCLVSIILYRAI--MAIVVSRSGNTLLAA-----W-----ASRIASL	547
		: : : : : : :: : : : : ::	
Db	157	VFMVFIASVISVTAFPVLCRILNELRLIKDRAGIVVLAAGIINDIMGWILLALSILSSA	216
Qy	548	TGSVNVLVFILILS----KIYVSLAHVLTRWEMHRT----QTKFEDAFTLKVFIFQFVNF	599
		: ::: : : : : : : :	
Db	217	EGSPVNTVYILLITFAWFLIYFFPLKYLLRWVLIRTHELDRSKPSPLATMCILFIMFISA	276
Qy	600	YSS-----PVYIAFFKGRFVGYPGNYHTLFGVRNEEC-----AAGGCLIELAQ-	642
		: : : : : :	
Db	277	YFTDIIGVHPIFGAFIAGLVVPRDDHYVVKLTERMEDIPNIVFIPIYFAVAGLNVDLTLL	336
Qy	643	-----ELLVIMVGKQVINNMQEVLIPLKLG--WWQKFRLRSKKRKAGASAGASQGP	691
		: : : : : : :	
Db	337	NEGRDWGYVFATIGIAIFTKIISG---TLTAKLTGLFWRE-----ATAAGV----	379
Qy	692	WEDDYELVPCEGLFD-EYLEMVLQFGFVT-----IFVAACPLAPLFALLNNWVEIRLDAR	745
		: : : : : : : :	
Db	380	-----LMSCKGIVEIVVLTVGLNAGIISRKIFGMFV-----LMALVSTFVTTPLTQL	426
Qy	746	KFVCEY----RRPVAERAQDIG	763
		: : : :	
Db	427	VYPDSYRDGVRKSLSTPAEDDG	448

RESULT 8


```
Qy      362 VV-----GTLVFLVGCFLVFSDIPTQELCGSKDSFEMCPLC 397
      |:                               | :: | :|:|:      :  : :| :
Db      279 VLKAFSPVYIYRYFKRGGDRWTSLGGIMLSITGIEALFADLSHFVPSAVQIAFTV---- 334

Qy      398 LDCPFWLLSSACALAQAGRLFDH-----GGTVFFSLFMALWAVLLLEYWKRKSATL 448
      : | ||: : | | ||                               |:|:: |:| : | :: |||
Db      335 IVFPCLLLAYSQGAAYIRRYPDHVADAFYRSIPGSVYWPMFIIATAAAIVASQATISATF 394

Qy      449 -----AYRWDCSDYEDTEERPRPQFAASA-----PMTAPN-----PITGEDEPYF 488
      | | | || : :: : | : :| | :| |
Db      395 SLVKQALAHGCF-----PRVKVHTSRKFLGQIYVPDINWILMILCIAVTAG----F 442

Qy      489 PERSRARRMLAGSVVIVVMVAVVVMCLVSIILYRAIMAIVVSRSGNTLLAAWASRIASLT 548
      :|: :|||::| :| | :|::| :| |
Db      443 KNQSQIGNAYGTAVVIVMLVTLLMTLIMILVWRCHWVLV-----LI 484

Qy      549 GSVVNLV-----FILILSKI-----YVSLAHVLTREWEMHRTQTKFEDAFTLKVFIFQ 595
      :|::|| | :| || :: | :| | | ||| :| :
Db      485 FTVLSLVVECTYFSAMLFKIDQGGWVPLVIAAFLIMVWVWHYG-----TLKRYEFE 536
```

RESULT 9
A45573
genome polyprotein - hepatitis C virus (strain JT)
N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstructural protein NS3); major envelope protein E; nonstructural protein NS1; nonstructural protein NS2; nonstructural protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
C;Species: hepatitis C virus
C;Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 09-Jul-2004
C;Accession: A45573
R;Tanaka, T.; Kato, N.; Nakagawa, M.; Ootsuyama, Y.; Cho, M.J.; Nakazawa, T.; Hijikata, M.; Ishimura, Y.; Shimotohno, K.
Virus Res. 23, 39-53, 1992
A;Title: Molecular cloning of hepatitis C virus genome from a single Japanese carrier: sequence variation within the same individual and among infected individuals.
A;Reference number: A45573; MUID:92295714; PMID:1318627
A;Accession: A45573
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-3010 <TAN>
A;Cross-references: UNIPROT:Q00269; UNIPARC:UPI0000131E29; GB:D11168; GB:D01171; NID:g221612; PIDN: BAA01943.1; PID:g221613
A;Experimental source: HCV-JT
A;Note: sequence extracted from NCBI backbone (NCBIN:106206, NCBIP:106207)
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: ATP; glycoprotein; hydrolase; nucleotide binding; P-loop; polyprotein; serine proteinase; transmembrane protein
F;2-115/Product: capsid protein C #status predicted <CPC>
F;116-191/Product: envelope protein M #status predicted <EPM>
F;192-389/Product: major envelope protein E #status predicted <MEE>
F;390-729/Product: nonstructural protein NS1 #status predicted <NS1>
F;730-1006/Product: nonstructural protein NS2 #status predicted <NS2>
F;1007-1615/Product: hepacivirin #status predicted <NS3>
F;1230-1237/Region: nucleotide-binding motif A (P-loop)
F;1312-1317/Region: nucleotide-binding motif B
F;1316-1319/Region: DEXH motif
F;1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>
F;1863-2013/Product: nonstructural protein NS4b #status predicted <N4B>
F;2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>

Query Match 2.2%; Score 108; DB 1; Length 3010;
Best Local Similarity 20.4%; Pred. No. 7.4;
Matches 160; Conservative 86; Mismatches 276; Indels 262; Gaps 41;

Qy	128	SAARDRTDMHRTWRETFLDNLRAAGLCVDQ-----QDVQDGNTTVHYALLSASWAVLCYY	182
Db	314	SGHRMAWDMMMNWSPT-----TALVVSQLLRIPQAVVDMVAGAHWGVL-----AGLAYY	362
Qy	183	AEDLRLKLPLQELPNQASNWSAGLLAWL---GIPNVLLEVVPDVPPEYYSCRFRVKNLPR	239
		: : : :	
Db	363	-----SMVGNWAKVLIVMLLFAGVDGVT-----YTTG-----	389
Qy	240	FLGSDNQDT-----FFTSTKRHQILFEILAKTPYGHEKKNLLGIHQLLAEGVLSAAFLPH	294
		: : : : : :	
Db	390	--GSQARHTQSVTSFFTQGPAQRI--QLINTNGSWHINRTALNCNESLNTGFFAALFYAH	445
Qy	295	DGPFKTPPE-----GP---QAPR-LNQRQVLFQHWA-----	321
		:	
Db	446	KFNSSGCPERMASCSSIDKFAQGWGPITYTEPRDLQRPYCW-HYAPRQC GIVPASQVCG	504
Qy	322	-----RWGK-----WNKYQPLDHVRRYFGEKVALYFAWL-	350
		: :	
Db	505	PVYCFTPSPVVVGTTDRSGAPTYNWGANETDVLLLNNT RPPQ--GNWFG-----CTWMN	556
Qy	351	--GFYTGWLLPAAVVG-----TLVFLVGCFLVFSDIPTQELCGSKDSFEMCPLCL-DCPF	402
		: : :	
Db	557	STGFTKTCGGPPCNIGGVGNLT LTCPTDCFRKHPE-ATYTKCGSGP--WLT PRCIVDYPY	613
Qy	403	WLLSSACALAQAGRLFDHGGTVFFSLFMALWAVLLLEYWKRKSATLAYRWDCSDYEDTEE	462
		: : : :	
Db	614	-----RLWHYPCTVNF TIFKVRMYVGGVEH--RLSA--ACNWTRGERCDLED	656
Qy	463	RPRPQ-----FAASAPMTAPNPITGEDEPYFPERSRARRMLAGSVVIVVMVAVVVMCLVS	517
		: : : : :	
Db	657	RDRSELSPLLLSTTEWQTLPCSFT-----TLPALSTGLIHLHQNIVDVQYLYGIGSAVVS	711
Qy	518	IILYRAIMAIVVSRSGNTLLAAWASRIASLTGSVNVLFILILSKIYVSLAHVLTRWEMH	577
		: : : : : : : : : : :	
Db	712	FVIKWEYIVLLF-----LLLADARVCACLW-----MMLLIAQAEAALENLVV--LN	755
Qy	578	RTQTKFEDAFTLVKFIFQFVNFYSSPVYIAFFKGRFVGYPGNYHTLFGV-----	626
		: : : : :	
Db	756	AASLAGADG-----ILSFLVFFCAAWYI---KGRLV--PGAAYALYGVWPLLLLLLLALP	804
Qy	627	-----RNEECAAGGCLIELAQELLVIMVG--KQVINNMQEV LIPKLKGWWQKFRLRSK	677
		: : : : : : : :	
Db	805	PRAYAMDREMAASCGG-----VVFVGLILLTLSPHYKVFLARLIWWLQYFITRAE	854
Qy	678	KRKAGASAGASQGPWEDDYELVPC----EGLFDEYLEMVLQFGFVTIFVAACPLAPLFAL	733
		: : : : : : :	
Db	855	AHLCVWVPPLNVRGGRDAIILLTCAAHPELIFDITKLLAILGLPLMVLQAAITAMPYFVR	914
Qy	734	LNNWVEIRLDARK-----FVCEYRRPVAERAQDIGIWFHILAGLT	773
		: : : : :	
Db	915	AQGLIRACMLVRKVAGGHYVQMAFMKLAALTGT YVYDHLTPL----QD---WAH--AGLR	965
Qy	774	HLAV 777	
Db	966	DLAV 969	

RESULT 10
T11129
cytochrome-c oxidase (EC 1.9.3.1) chain I - acorn worm mitochondrion
C;Species: mitochondrion Balanoglossus carnosus
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004

Query Match 2.2%; Score 106.5; DB 2; Length 519;
Best Local Similarity 17.8%; Pred. No. 0.93;
Matches 94; Conservative 51; Mismatches 167; Indels 215; Gaps 18;

Qy	472	APMTAPNPITGEDEPYFPERSRARRMLAGSVVIVVMVAVVVMCLVSIILYRAIMAVVSR	531
		: : : : : : : :	
Db	39	AELAQPGLLGDDQIY-----NVIVTAHAFVMIFFMVMPIMIGG	77
Qy	532	SGNTLL-----AAW-----ASRIA	545
Db	78	FGNWLLPLMLGAPDMAFPRLNMSFWLLPPSFLLLLSSAGVESGVGTGWTVYPPLAGNMA	137
Qy	546	SLTGSVVNLVFILIL---SKIYVSLAHVLTRWEMHRTQTKFE--DAFTLKVFIFQFVNFY	600
		: : : : : :	
Db	138	HAGGSVDLAIFSLHLAGISSILGAINFMTTVINMRAPGVRFDRLPLFVWSVFITVILLLL	197
Qy	601	SSPV-----Y	605
Db	198	SLPVLAGAITMLLTDRNLNTSFFDPAGGGDPILYQHLFWFFGHPEVYILILPAFGMISHV	257
Qy	606	IAFFKGRF--VGYPGNYHTLFGVRNEECAAGGCLIELAQELLVIMVGKQVINN-----M	657
		: : : : :	
Db	258	IAFYSGKKEPFGYLGMYAMIAI-----GILGFLVWAHHMFTVGMDVDTRAYFTAAT	309
Qy	658	QEVLP---KLKGWWQKFRLRSKKRKAGASAGASQGPWEDDYELVPCEGLFDEYLEMVLO	714
		: : : : : :	
Db	310	MVIAVPTGIKIFSWL-----ATLHGSALQWE-----APLLWA	341
Qy	715	FGFVTIFVAACPLAPLFAALLNNWVEIRLDARKFVCEYRRPVAERAQDIGI-----WFHI	768
		: : : : : : :	
Db	342	LGFVFLFTVGGLTG--IVLSNSSLDVVMHDTYYVVAHFHYVLSMGAVFGIFAGFIHWFPL	399
Qy	769	LAGLTHLAVISNAFLLAFSSDFLPRAYYRWTRAHDLRGFLNFTLARAPSSFAAAHNRTC	828
		:	
Db	400	FTGLTLHPV-----WTKFHFWTMFLGVNLTFFPQHFLGLAGMPRR	439
Qy	829	YRAFRDDDDGHYSQTYWNLLA-----IRLAFVIVFEHVV---FSVGRL	867
		: : : : : :	

Db 440 YSDYPD-----AYTTWNVLSSVGSIVSLASVIIFLAILWEAFTARRL 481

RESULT 11

B86088

probable citrate permease Z5523 [imported] - Escherichia coli (strain O157:H7, substrain EDL933)

C;Species: Escherichia coli

C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004

C;Accession: B86088

R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew, G.F.; Evans, P.S.; Gregor, J.; Kirkpatrick, H.A.; Posfai, G.; Hackett, J.; Klink, S.; Boutin, A.; Shao, Y.; Miller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca, J.; Anantharaman, T.S.; Lin, J.; Yen, G.; Schwartz, D.C.; Welch, R.A.; Blattner, F.R.

Nature 409, 529-533, 2001

A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.

A;Reference number: A85480; MUID:21074935; PMID:11206551

A;Accession: B86088

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-438 <STO>

A;Cross-references: UNIPROT:Q8X4P7; UNIPARC:UPI00000D0D9E; GB:AE005174; NID:g12518889; PIDN:

AAG59166.1; GSPDB:GN00145; UWGP:Z5523

A;Experimental source: strain O157:H7, substrain EDL933

C;Genetics:

A;Gene: Z5523

C;Superfamily: citrate utilization determinant

Query Match 2.1%; Score 106; DB 2; Length 438;
Best Local Similarity 23.3%; Pred. No. 0.82;
Matches 57; Conservative 31; Mismatches 69; Indels 88; Gaps 13;

Qy 339 FGEKVAL-YFAWLGFYT-----GWLLPAAVVGTLVFLVGCFLVFS---DIP---T 381
|| ||| ||| | | || :| :| |: ||:| | |:|

Db 173 FGGVVALGLSAWLPFATGSETVMAEWGWRVP-FFIGVLLAPVGCWLRLSLENDVPEPAHN 231

Qy 382 QELCGSKDSFEMCPLCLDCPFWLLSSACALAQAGRLFDHGGTV--FFSLF-MALWAVLLL 438
:: |: :| : |: :| | | | || : ||| || |

Db 232 KKAAASESAFSL-----LMQHKATIAN-GILLAIGSTVATYISLFYYGTWAAKYL 280

Qy 439 EYWKRSATLAYRWDCSDYEDTEERPRPQFAASAPMTAPNPITGEDEPYFPERSRARRML 498
| :: | | | :|

Db 281 -----GMNQNY----SHAAMLL 293

Qy 499 AGSVVIVVMVAVVVMC----LVSIILYRAIMAIVVSRSGNTLLAAWASRIASLTGSVVNL 554
|| : | : | ::| :|| :| ::| || :| | :::

Db 294 AGVITFVGALLVGMLCDSVGRKKLILISRMVVLICSWPSFWLLVNYPS-----PGMLLTV 348

Qy 555 VFILI 559
||:::

Db 349 VFVMV 353

RESULT 12

E91240

probable membrane transport / symporter protein ECs4893 [imported] - Escherichia coli (strain O157:H7, substrain RIMD 0509952)

C;Species: Escherichia coli

C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004

C;Accession: E91240

R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.; Ohtsubo, E.; Nakayama, K.; Murata, T.; Tanaka, M.; Tobe, T.; Iida, T.; Takami, H.; Honda, T.; Sasakawa, C.; Ogasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.

DNA Res. 8, 11-22, 2001

A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genomic comparison with a laboratory strain K-12.
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: E91240
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-438 <HAY>
A;Cross-references: UNIPROT:Q8X4P7; UNIPARC:UPI00000D0D9E; GB:BA000007; PIDN:BAB38316.1; PID:g13364369; GSPDB:GN00154
A;Experimental source: strain O157:H7, substrain RIMD 0509952
C;Genetics:
A;Gene: ECs4893
C;Superfamily: citrate utilization determinant

Query Match	2.1%;	Score 106;	DB 2;	Length 438;
Best Local Similarity	23.3%;	Pred. No. 0.82;		
Matches	57;	Conservative 31;	Mismatches 69;	Indels 88; Gaps 13;

Qy	339	FGEKVAL-YFAWLGFYT-----GWLLPAAVVGTLVFLVGCFLVFS---DIP---T	381
		: : : :	
Db	173	FGGVVALGLSAWLPFATGSETVMAEWGWRVP-FFIGVLLAPVGCWLRLSLENDVPEPAHN	231

Qy	382	QELCGSKDSFEMCPLCLDCPFWLLSSACALAQAGRLFDHGGTV--FFSLF-MALWAVLLL	438
		:: : : : : : :	
Db	232	KKAAASESAFSL-----LMQHKATIAN-GILLAIGSTVATYISLFYYGTWAAKYL	280

Qy	439	EYWKRK SATLAYRWDCSDYEDTEERPRPQFAASAPMTAPNPITGEDEPYFPERSRARRML	498
		: : :	
Db	281	-----GMNQNY----SHAAMLL	293

Qy	499	AGSVVIVVMVAVVVMC----LVSIILYRAIMAIVVSRSGNTLLAAWASRIASLTGSVVNL	554
		: : :: : : :: : :: :	
Db	294	AGVITFVGALLVGMLCDSVGRKKLILISRVMVLICSWPSFWLLVNYP-----PGMLLTV	348

Qy	555	VFILI	559
		:::	
Db	349	VFVMV	353

RESULT 13
JC1346
dopamine beta-monooxygenase (EC 1.14.17.1) precursor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
C;Accession: JC1346
R;Nakano, T.; Kobayashi, K.; Saito, S.; Fujita, K.; Nagatsu, T.
Biochem. Biophys. Res. Commun. 189, 590-599, 1992
A;Title: Mouse dopamine beta-hydroxylase: primary structure deduced from the cDNA sequence and exon/intron organization of the gene.
A;Reference number: JC1346; MUID:93080618; PMID:1280432
A;Accession: JC1346
A;Molecule type: mRNA
A;Residues: 1-621 <NAK>
A;Cross-references: UNIPROT:Q64237; UNIPARC:UPI0000029950; GB:S50200; NID:g260872; PIDN:AAB24330.1; PID:g260873
C;Comment: This enzyme catalyzes the hydroxylation of dopamine to norepinephrine.
C;Genetics:
A;Introns: 117/3; 166/3; 252/3; 346/1; 401/3; 449/3; 462/3; 482/3; 525/2; 578/3
C;Keywords: catecholamine biosynthesis; copper; glycoprotein; monooxygenase; oxidoreductase; phosphoprotein
F;1-43/Domain: (or 1-46) signal sequence #status predicted <SIG>
F;44-621/Product: (or 47-621) dopamine beta-monooxygenase #status predicted <MAT>
F;300-523/Domain: peptidylglycine monooxygenase I homology <PGM>

F;68,188,476,570/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;350,528/Binding site: phosphate (Ser) (covalent) (by calmodulin-dependent kinase II) #status predicted

Query Match 2.1%; Score 105; DB 2; Length 621;
Best Local Similarity 20.2%; Pred. No. 1.6;
Matches 113; Conservative 61; Mismatches 187; Indels 198; Gaps 29;

Qy	110	ADFVLVWEE-----DLKLDRQQD---SAARDRTDMHRTWRETFLDNLRA	150
		: : : : : : : :	
Db	102	ADLIMLWSDGDRAYFADAWSDRKGQIHLSQQQDYQLLQAQRTDGLSLLFKRPF-----	155
Qy	151	AGLCVDQQDVQDGNTTVH--YALLSASWAVLCYYAEDLRLKLPLQELPNQASNWSAGLLA	208
		: : : : : :	
Db	156	--VTCDPKDYVIEDDTVHLVYGILEE-----PFQSL--EAINTS-----	190
Qy	209	WLGIPNVLLEV-----VPDVPPEYYSCRFRVKNLPRFLGSDNQDTFF-----TS	252
		: : : : : : :	
Db	191	--GLHTGLLRVQLLKSEVPTSPMPEDVQTMDIRA---PDILIPDNEQTYWCYITELPPRF	245
Qy	253	TKRHQILFEILAKTPYGHEKKNLLGIHQLLAEGVLSAAFPPLHDGP--FKTPPEGPQAPRL	310
		: : : : : : : : :	
Db	246	PRHHIIMYEAIIV-TEGNEALVHHMEVFQCAAE---SEDFPQFNGPCDSKMKPD-----RL	296
Qy	311	NQRQVLFQHWARWGKWNKYQPLD-----HVERRYFGEK---VAL	345
		: : : : : : : :	
Db	297	NYCRHVLAAWALGAK-AFYYPKEAGVPFGGPGSSPFLRLEVHYHNPRKIQGRQDSSGIRL	355
Qy	346	-YFAWLGFYTGWLLPAAVVGTLVFLVGCFLVFSDIPTQELCGSKDSFEMCPLCLD-CPFW	403
		: : : : : : :	
Db	356	PYTATLRRYDAGIMELGLVYTPLMA-----IPPQE-----TAFVLTGYCTDKCTQM	401
Qy	404	LLSSACALAQAGRLFDHGGTVFFSLFMALWAVLLLEYWKRKSATLAYRWDCSDYEDTEER	463
		: : : : : :	
Db	402	ALQDSGIHIFASQLHTH-----LTGRKVVTVLAR-----DGQER	435
Qy	464	PRPQFAASAPMTAPNPITGEDEPYFPERSRARRMLAGSVVIVVMVAVVVMCLVSIILYRA	523
		: : :	
Db	436	KE-----VNRDNHYSF-HFREIRMLKKVVTVYPGDVLITSC-----	470
Qy	524	IMAIVVSRSGNTLLAAWASRIASLTGSVNVLVFILILSKIYVSLAHVLTRWEMHRTQTKF	583
		: : : : : : : : :	
Db	471	----TYNTENKTL-----ATVGG-----FGILEEMCVNYVHYYPQTELELCKSAV	511
Qy	584	EDAFTLKVFIQFVNFYSS	602
		: :	
Db	512	DDGFLQK--YFHMVNRFS	528

RESULT 14
H82555
c-type cytochrome biogenesis membrane protein XF2460 [imported] - Xylella fastidiosa (strain 9a5c)
C;Species: Xylella fastidiosa
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C;Accession: H82555
R;anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequencing and Analysis, Sao Paulo, Brazil.
Nature 406, 151-157, 2000
A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A;Reference number: A82515; MUID:20365717; PMID:10910347
A;Note: for a complete list of authors see reference number A59328 below
A;Accession: H82555
A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-646 <SIM>

A;Cross-references: UNIPROT:Q9PAN5; UNIPARC:UPI00000C2A60; GB:AE004054; GB:AE003849; NID:g9107645; PIDN:AAF85259.1; GSPDB:GN00128; XFSC:XF2460

A;Experimental source: strain 9a5c

R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; Alves, L.M. C.; Araya, J.E.; Baia, G.S.; Baptista, C.S.; Barros, M.H.; Bonaccorsi, E.D.; Bordin, S.; Bove, J.M.; Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H.; Colauto, N.B.; Colombo, C.; Costa, F.F.; Costa, M.C.R.; Costa-Neto, C.M.; Coutinho, L.L.; Cristofani, M.; Dias-Neto, E.; Docena, C.; El-Dorrry, H.; Facincani, A.P.; Ferreira, A.J.S.

submitted to GenBank, June 2000

A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohme, M.; Furlan, L.R.; Garnier, M.; Goldman, G.H.; Goldman, M.H.S.; Gomes, S.L.; Gruber, A.; Ho, P.L.; Hoheisel, J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigret, F.; Lambais, M.R.; Leite, L.C.C.; Lemos, E.G.M.; Lemos, M.V.F.; Lopes, S.A.; Lopes, C.R.; Machado, J.A.; Machado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E.A.L.

A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.; Monteiro-Vitorello, C.B.; Moon, D.H.; Nagai, M.A.; Nascimento, A.L.T.O.; Netto, L.E.S.; Nhani Jr., A.; Nobrega, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A.; Paris, A.; Peixoto, B.R.; Pereira, G.A.G.; Pereira Jr., H.A.; Pesquero, J.B.; Quaggio, R.B.; Roberto, P.G.; Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasaki, H.E.

A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveira, J.F.; Silvestri, M.L.Z.; Siqueira, W.J.; de Souza, A.A.; de Souza, A.P.; Terenzi, M.F.; Truffi, D.; Tsai, S.M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Zago, M.A.; Zatz, M.; Meidanis, J.; Setubal, J.C.

A;Reference number: A59328

A;Contents: annotation

C;Genetics:

A;Gene: XF2460

C;Superfamily: nrFE protein

Query Match 2.1%; Score 104; DB 2; Length 646;
Best Local Similarity 18.9%; Pred. No. 2;
Matches 95; Conservative 46; Mismatches 152; Indels 210; Gaps 19;

```

Qy      167 VHYALLSASWAVLC-----YYAEDLRLKLPLQELPNQASNWSA---GLLAWLGI 212
      |  :||: ::|:|      | ||:  ||  |  : ||  || |
Db      45 VQLSLLAGAFALLTYAFLGNDFSVQYVAENSHSLP--TLYRSTAVWGAHEGSLLLW--- 99

Qy      213 PNVLLEVPDVPPEYYSCRFRVKNLPRFLGSDNQDTFFTSTKRHQILFEILAKTPYGHEK 272
      |||      :      |  | : ||  | :
Db      100 --VLL-----LAGWTASVALRSHTLPATLSA----- 123

Qy      273 KNLLGIHQLLAEGVLSAAFPLHDGPF----KTPPEGQPAPRLNQRQVLFQH----- 319
      :||:  |:| || |  ||      |||      ||  :  |
Db      124 -RILGVLGLIALGFL-ALILFTSNPFARLLPAVPEGNDLNPLLQDPGMIVHPPLLYAGYI 181

Qy      320 -----WARW-----GKWNKYQPLDHVRRYFG 340
      | ||      | || | |  | :|
Db      182 GFAVPFAFAVAVLLEGRIDPTWLRWSRPWHTAWALLTLGIALGSWWAYYELGWGGWWFW 241

Qy      341 EKV--ALYFAWL-----GFYTGWLLPAAVVGTLVFLVGCFLVFSDIPTQE 383
      : |  | : ||      | : | | | :  : | : |
Db      242 DPVENASFMPWLIGVALIHSQAITDKRGSFTHWTLTLLAITAFALALLGTFLVRSSVLT-- 299

Qy      384 LCGSKDSFEMCPLCLDCPFWLLSSACALAQAGRLFDHGGTVFFSLFMALWAVLLLEYWKR 443
      |  :|  |:      |  :  :|  :  |||
Db      300 ---SVHAFAADPV-----RGAFILLLIFTLIGGALLL----- 328

Qy      444 KSATLAYRWDCSDYEDTEERPRPQFAASAPMTAPNPITGEDEPYFPERSRARRMLAGSVV 503
      :|  ||      |:|  :  |  ||  :|  ::

```

Db 329 -----YARRAPQL--TPVTINMQQRFTPVSRETLLLLNNLL 362

Qy 504 IVVMVAVVVMCLVSIILYRAIMAIIVVSRSGNTLLAAWASRIA----SLTGSVVNLVFILI 559
: |:|:: || | :| |: :| |:

Db 363 LTCACAMVLL-----GTLYPLLADALALGQLSVGPPYFGPLFTLL 402

Qy 560 LSKIIYVSL-AHVLTRWEMHRTQT 581
:: : || |||: |

Db 403 MTPLIVLLPLGPFTRWQREHPST 425

RESULT 15

JQ2034

RNA-directed RNA polymerase (EC 2.7.7.48) - beet cryptic virus 3

C;Species: beet cryptic virus 3

C;Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 09-Jul-2004

C;Accession: JQ2034

R;Xie, W.S.; Antoniw, J.F.; White, R.F.

J. Gen. Virol. 74, 1467-1470, 1993

A;Title: Nucleotide sequence of beet cryptic virus 3 dsRNA2 which encodes a putative RNA-dependent RNA polymerase.

A;Reference number: JQ2034; MUID:93329401; PMID:8336129

A;Accession: JQ2034

A;Molecule type: mRNA

A;Residues: 1-478 <XIE>

A;Cross-references: UNIPROT:Q86632; UNIPARC:UPI00000EE5F9; GB:S63913; NID:g407557; PIDN:AAB27624.1; PID:g407558

C;Genetics:

A;Map position: segment RNA2

C;Keywords: nucleotidyltransferase; reverse transcriptase

Query Match 2.1%; Score 103.5; DB 2; Length 478;
Best Local Similarity 20.7%; Pred. No. 1.5;
Matches 53; Conservative 34; Mismatches 96; Indels 73; Gaps 11;

Qy 59 RAQEEDSTVLIDVSPPEAEKRGSYGSTAHASEPGGQQAACRAGSPAKPRIADFVLVWEE 118
:|: | :| || || :| |||: || | || : : ||

Db 109 KARAFDVNTELDKVPYEQSSSAGYGYRSHKGPPEGGE--THMRAISRVKPTLMTAIRPDEE 166

Qy 119 -----DLKLDRQQDSAARDRTDMHRTWRETFLDNLRAAGLCVDQQDVQDGNTTV 167
|: | | : |:| : | |

Db 167 GPEYTILESVPDIGYTRTQLADLREKTKVRGVWGRAF----- 203

Qy 168 HYALLSASWA-----VLCYYAEDLRLKLP--LQELPNQASNWSAGLLAWLGIPN 214
|| |: : | : | :| :| :| :| :| :| :| :|

Db 204 HYILIEGTAARPLENFMLGTTFMHIGSDPQLSVPRILHQMKREGSKWLYA-LDWSSFDS 262

Qy 215 VLLEVVPDPVPEYYSCRFRVNKLPRFLGSDNQDTFFTSTKRHQILFEILAKTPYGHEK-- 272
| :| | | | : |:| :| | |:| : |:|

Db 263 -----SVTRFEINCAF--NLLKERIEFPNEET-----ELAFE-LSRILFKHKKLA 304

Qy 273 ---KNLLGIHQLLAEG 285
|: ||: : |

Db 305 APDGNIYMIHKGIPSG 320

Search completed: June 24, 2008, 08:37:14

Job time : 37 secs